

Charles University in Prague
Faculty of Science
Department of Physical and Macromolecular
Chemistry



Interaction of heteroboranes with biomolecules:
Accurate quantum chemical study

Doctoral Thesis Abstract

RNDr. Jindřich Fanfrlík

Advisor: Prof. Ing. Pavel Hobza, DrSc.

Institute of Organic Chemistry and Biochemistry, AS CR
Center for Biomolecules and Complex Molecular Systems

Prague 2008

Univerzita Karlova v Praze
Přírodovědecká Fakulta
Katedra fyzikální a makromolekulární chemie



Interakce heteroboranů s biomolekulami:
kvantově chemická studie

Autoreferát disertační práce

RNDr. Jindřich Fanfrlík

Školitel: Prof. Ing. Pavel Hobza, DrSc.

Ústav organické chemie and biochemie, AV ČR
Centrum biomolekul a komplexních molekulových systémů

Praha 2008

Introduction

The world wide epidemic of AIDS is caused by two species of the human immunodeficiency virus (HIV-1 and HIV-2). Its 9200-bases-long RNA genome contains three main genes: *gag*, *pol* and *env*. The enzymes located within the *pol* gene (protease PR and reverse transcriptase RT) have become major targets for drug discovery, e.g. RT inhibitor Tenofovir disoproxil fumarate (DF) discovered by A. Holy at UOCHB in Prague and manufactured by Gilead Sciences and also already ten HIV-1 PR inhibitors (PIs) have been approved for use in the clinics:

The evolution of drug resistant variants represents a major complication of highly active anti-retroviral therapy. Due to the rapid selection of resistant viral strains, the development of new, effective inhibitors with high genetic barrier to resistance development is still very relevant. Recently, Petr Cigler with coworkers in the laboratory of Jan Konvalinka has discovered a group of novel, potent, specific and resistant PIs; the polyhedral metallacarborane compounds (MCB). Ionic metal bis(dicarbolides) consists of two dicarbolide subclusters sandwiching the central metal atom. The equal 11-vertex dicarbolide subclusters are, thus, covalently connected by a *commo* metal vertex, forming two 12-vertex metal dicarbolide subclusters. The skeletal structure of the parental compound of the metallacarborane PIs, $[3,3' \text{-Co}(1,2\text{-C}_2\text{B}_9\text{H}_{11})_2]^{-1}$ designated as COSAN in the following text. Connecting of two parent cages with a linker results in compounds with significantly increased inhibitory potency toward HIV PR. ¹ Such a compound is designed GB80 in the following text.

The aim of this work is to study the interactions of metallacarborane inhibitors with their biological partner. These unique interactions can be described reliably only using high-level quantum chemistry methods, so we start with the simplest model of metallacarboranes inhibitors: 1-carba-*closo*-dodecaborane ($\text{CB}_{11}\text{H}_{12}$). We employ $(\text{CB}_{11}\text{H}_{12})^-$ and study its interaction with building blocks of biomolecules in order to unveil the nature of carborane-biomolecule interactions.

In the next step, we study the effect of various metal vertices and exo-substitutions on several series of metallacarboranes. We focus on their interactions with a tetrapeptide, which model their biomolecular partner. We compare the effect of the

substitutions on the strength of stabilisation energies and on the height of desolvation penalties.

In the third part of the thesis, we examine the X-ray structure of the HIV PR-COSAN complex by using hybrid QM/MM approach and aim at refining atomistic details of the structure, which are crucial for understanding of energetic of the inhibition of HIV PR by MCBs. We begin by deciding on the identity of the residues in Positions 50 and 128, whether these are water molecules or Na^+ counterions. Na^+ could serve as a counterion to the negatively charged molecules of GB18 in a similar way as it does in metallacarborane aggregates in solution. Subsequently, the energy profiles for the rotation of Cb1 and Cb2 molecules in the PR cavity are calculated. The combination of these two energy profiles along with the evaluation of their inter-dependence yields energetically favorable and unfavorable positions for carbon atoms in GB18 complexed to PR. It should be mention here that the positions of boron and carbon atoms of COSAN bound in the PR active-site are indistinguishable by X-ray crystallography as well as Na^+ ions and water molecules.

In the fourth part of the thesis, the effect of active-site mutations Val82Thr and Ile84Val of HIV-1 PR on the binding of COSAN was studied. We explain the small 6- and 2-fold decrease in binding strength of COSAN for HIV PR mutants.

In the fifth part of this thesis, we study crystal structure of wild-type HIV-1 PR inhibited by compound GB-80. No continuous map for model linker connecting two cobaltacarborane cages of GB-80 occupying HIV PR dimer was experimentally observed and, thus, possible and energetically favorable linker conformations in the enzyme cavity are calculated

Methods

Due to the complexity of the potential energy surfaces of molecular complexes, we have used the molecular dynamics (MD) simulations combined with the quenching (Q) technique to explore PESs of given complexes. MD/Q simulations were carried out using two different empirical potentials. First, we employed Universal Force Field (UFF) potential within the TINK program to explore the PES of the complexes of $(\text{CB}_{11}\text{H}_{12})^-$

with building block of biomolecules (*i.e.* no water was considered). Second, we used AMBER8 package to obtain possible conformers of GB80 in the GB80-PR complex (continuous water model was considered).

For the gradient optimisation, we used the DFT method augmented by empirically parameterised dispersion correction (D). We applied the resolution of the identity (RI) approximation to the DFT method combined with the TPSS functional and SVP (3s2p1d/2s1p) basis set. Interaction energies of all the studied complexes were determined using the RI-DFT-D methodology at the TPSS/TZVP level. These calculations were performed using Turbomole 5.8.

Hydration free energies were calculating using the COSMO (C-PCM) implicit solvent model implemented in the Gaussian03 code. The recommended HF/6-31G* level combined with the united atom radii (UAHF) model was used.

We performed the QM/MM calculation by our own QM/MM code. It acts as an interface between Turbomole package ² that is called to do the QM calculations as well as geometry optimizations, our code used to calculate the dispersion correction and AMBER package used for MM calculations. Sander module of AMBER 7 package was modified to write out gradients.

Results

In the first part, we found that 1-carba-*closo*-dodecaborane ($\text{CB}_{11}\text{H}_{12}$), a model for metallacarborane inhibitors of HIV protease, interacts with building blocks of biomolecules via formation of dihydrogen bonds and no other binding motif was observed. Dihydrogen bonds are mainly electrostatic interactions between negatively-charged boron-bound hydrogen atoms and positively-charged hydrogen atoms of biomolecules. We find that carboranes forms dihydrogen bonds with biomolecules preferably with atoms of its lower hemisphere, *i.e.* the part of the cage opposite to the carbon atom.

In the second part, we showed that the total negative charge of heteroboranes is located mainly on boron-bound hydrogens. An increase of negative total charge (from 0 to -1 or -2) of heteroboranes yields an increase in stabilisation energies of heteroborane... peptide complexes and also a huge increase in hydration free energies of heteroboranes.

Since these two energy terms compensate, both must be taken into account when studying the interactions of carboranes with biomolecules. When the exo-substitutions of metallacarboranes are compared to the substitutions of metal vertices, they cause a higher increase in stabilisation energies and a lower increase in desolvation penalties.

In the third part, the binding of COSAN to HIV-1 protease was interpreted in atomistic and energy details using high-level QM/MM calculations. The rotation profiles of the individual COSAN molecules in PR surroundings were evaluated. A large effect of sodium counterions was detected through its being energetically disfavored for the closest boron atom positions to be occupied by carbon atoms. Subsequently, the interdependence of the rotational profiles was assessed. Based on these results, the favorable positions for carbon atoms in the dicarbollide cages of COSAN bound to HIV-1 PR were extrapolated and delineated. Specifically, in the independent Cb2 molecule, three rotamers are supposed to be well populated (within 1 kcal/mol) and another six, less stable rotamers could also be present (within 5 kcal/mol). The Cb1 molecule is dependent on the rotation of Cb2, thus making it impossible for all the combinations (in the order of hundreds) to be determined and allowing for an exclusion of seven rotamers of Cb1 out of twenty-five only by estimation.

In the fourth part, the effect of active-site mutations Val82Thr and Ile84Val of HIV-1 PR on the binding of COSAN was studied. The small 6- and 2-fold decrease in binding strength of COSAN for HIV PR mutants is caused by the loss of only one dihydrogen bond according to our model.

In the fifth part, we modeled possible and energetically favorable conformations of the linker connecting two cobaltacarborane cages of **GB80** in the enzyme cavity. As a result, we obtained a series of conformations of GB80 in PR with a wide range of energies. Altogether, five conformers of the GB80 linker were in a chosen range of 3kcal/mol in the simulation. These five GB80 structures differ in the conformations of the linker, while the position of the central nitrogen atom is conserved and corresponds to the electron density into which water molecule number 110 is fitted in the X-ray structure.

Conclusions

i) Carboranes interact with building blocks of biomolecules via formation of dihydrogen bonds.

ii) Carboranes forms dihydrogen bonds with biomolecules preferably with atoms that are located on the part of the cage cage opposite to the carbon atom.

iii) The exo-substitutions of metallacarboranes cause a higher increase in stabilisation energies and a lower increase in desolvation penalties compared to the substitutions of metal vertices.

iv) Three rotamers of Cb2 are predicted to be well populated (within 1 kcal/mol) and another six, less stable rotamers could also be present (within 5 kcal/mol).

v) The Cb1 molecule is dependent on the rotation of Cb2, thus making it impossible for all the combinations (in the order of hundreds) to be determined and allowing for an exclusion of seven rotamers of Cb1 out of twenty-five only by estimation.

vi) The small 6- and 2-fold decrease in binding strength of COSAN for mutations Val82Thr and Ile84Val of HIV-1 PR is caused by the loss of only one dihydrogen bond according to our model

vii) A series of conformations of GB80 in PR with a wide range of energies was obtained. These GB80 structures differ in the conformations of the linker, while the position of the central nitrogen atom is conserved and corresponds to the electron density into which water molecule number 110 is fitted in the X-ray structure.

Úvod

Epidemie AIDS je způsobena 2 variantami viru anglicky pojmenovaného Human Immunodeficiency Virus (HIV-1 and HIV-2). Jeho RNA genom se skládá ze 3 hlavních genů: *gag*, *pol* a *env*. Enzymy lokalizované na *pol* genu (proteáza PR a reverzní transkriptáza RT) se staly hlavním cílem při hledání nových léčiv, např. RT inhibitor Tenofovir disoproxil fumarát (DF) objevený A. Holým na UOCHB AV ČR v Praze a zpracovaný firmou Gilead Sciences; již bylo schváleno 10 HIV-1 PR inhibitorů (PIs).

Rychlý vývoj rezistentních variant viru představuje největší překážku pro vysoce účinnou terapii, proto je nezbytný vývoj nových účinných inhibitorů odolných vůči rezistenci. Petr Cígler ve skupině Jana Konvalinky objevil skupinu nových, specifických a rezistentních inhibitorů HIV PR (PIs), kterými jsou polyhedrální metalokarborány (MCB). Iontové metalobisdikarbolidy se skládají ze dvou stejných 11-vrcholových karbolidových klecí obklopující středový atom, se kterým jsou kovalentně spojeny. Výchozí sloučenina metalokarboranových PIs, $[3,3'\text{-Co}(1,2\text{-C}_2\text{B}_9\text{H}_{11})_2]^{-1}$ je v textu značena jako COSAN. Spojení 2 COSANů flexibilním linkrem (sloučenina GB80) znatelně zvyšuje účinnost inhibitoru.

Cílem této práce je studium interakcí metalokarboranových inhibitorů s jejich biologickým partnerem. Tyto unikátní interakce mohou být řádně popsány jen za použití adekvátní kvantově chemické metody. Proto jsme naši studii začali na nejmenším možném modelu PIs; na *closo* dodekaboránu ($\text{CB}_{11}\text{H}_{12}$) a analyzovali jsme jeho interakce se stavebními bloky biomolekul.

V dalším kroku jsme studovali vliv záměny přechodného kovu a také exo-substitucí (např. B-OH nebo B-NH₂ za B-H). Zaměřili jsme se na interakce s tetrapeptidem, který modeloval jejich biomolekulárního partnera, a porovnávali jsme vliv modifikací metalokarboránů na sílu interakce a na změnu desolvatační energie.

Ve třetí části jsme studovali krystalografickou strukturu komplexu HIV PR-COSAN za použití kombinace kvantové a molekulové mechaniky (QM/MM metoda) a upřesňovali strukturu na atomární úrovni. Nejdříve jsme určovali, zdali je na pozici 50 a 128 přítomná voda nebo sodný ion. Na⁺ ion by zde mohl mít obdobnou úlohu jako při agregaci metalokarboránů v roztoku. Následně jsme počítali rotační profily COSANů (značených Cb1 a Cb2) v dutině HIV PR. Kombinací vypočtených rotačních profilů s

posouzením jejich vzájemné závislosti jsme určili, které rotamery COSANu mohou být přítomny v komplexu s PR. Bor je nerozlišitelný od uhlíku v X-ray krystalografii, a tak lze přítomné rotamery určit jen pomocí výpočetních metod.

Ve čtvrté části práce byl analyzován malý pokles síly vazby mezi COSANEM a mutovanou HIV PR (mutace Val82Thr a Ile84Val).

V poslední části práce byl studován komplex HIV-1 PR a jejího inhibitoru GB80. Hledali jsme konformace řetězce spojující kobaltokarboránové klece, jelikož jejich mapy elektronové hustoty nejsou naměřitelné.

Metody

Molekulová dynamika v kombinaci s “quenchingem (MD/Q)” byla použita k prohledávání povrchů potenciálních energií studovaných komplexů. MD/Q simulace byly provedeny za použití dvou různých empirických potencialů: 1) Universal Force Field (UFF) implementovaný do TINK programu při studování komplexů $(CB_{11}H_{12})^-$, 2) AMBER8 při hledání konformerů GB80 v komplexu s HIV PR.

Pro optimalizaci byla používána metoda DFT rozšířená o empirickou disperzi (D). Použili jsme “Resolution of the Identity” (RI) aproximaci v DFT metodě v kombinaci s TPSS funkcí a SVP (3s2p1d/2s1p) bázovým setem. Interakční energie byly určeny pomocí RI-DFT-D/TPSS/TZVP. Tyto výpočty byly provedeny v Turbomolu 5.8.

Volné hydratační energie byly určeny pomocí COSMO (C-PCM) implicitního modelu rozpouštědla implementovaného v Gausiánu03.

Pro QM/MM výpočty jsme použili QM/MM kód vyvinutý v naší laboratoři, jež spojuje Turbomol (QM výpočty) s naším kódem pro určení disperzní energie a AMBER (MM výpočty). Sander z AMBERu 7 byl použit k zápisu gradientů.

Výsledky

V první části této práce jsme zjistili, že $(CB_{11}H_{12})^-$ model metalokarboránových inhibitorů HIV PR interaguje se stavebními bloky biomolekul téměř výlučně tvorbou dvojvodíkových vazeb. Dvojvodíkové vazby jsou interakce mezi zápornými vodíky

karboránů a kladnými vodíky biomolekul. Dále jsme zjistili, že dvojevodíkové vazby jsou přednostně tvořeny vodíky v *meta* a *para* poloze.

V druhé části jsme ukázali, že záporný náboj je na karboránech lokalizován na vodících, jež jsou navázány na bóry. Vzrůst celkového záporného náboje (z 0 na -1 nebo -2) u heteroboránů zvyšuje stabilizační energii heteroboránového komplexu s peptidem, ale také výrazně zvyšuje desolvatační energii heterokarboránů. Jelikož se tyto dva příspěvky kompenzují, tak oba dva musí být brány v úvahu při studiu interakcí karboránů s biomolekulami. Exo-substituce metalokarboránů způsobují vyšší narůst stabilizačních energií a nižší narůst desolvatačních energií než substituce přechodného kovu.

Ve třetí části jsme analyzovali komplex COSANu a HIV-1 PR na atomární úrovni použitím QM/MM výpočtů a vyčíslili jsme rotační profily jednotlivých molekul COSANu v PR. Pozorovali jsme výrazný vliv sodných protiiontů. Komplexy PR s COSANy, jež obsahovaly takové rotamery, u nichž byly CH skupiny blízko sodných protiiontů, byly výrazně méně stabilní. Nasledně jsme našli nezávislost rotačního profilu Cb2. Na základě těchto výsledků jsme určili nejvýhodnější pozice pro uhlíkaté atomy karboránů v komplexu s PR. Konkrétně u více nezávislé Cb2 molekuly jsou tři rotamery silně populovány a dalších šest rotamerů může být také přítomných. Cb1 molekula je více závislá na rotaci Cb2 a sedm z pětadvaceti rotamerů Cb1 může být zcela vyloučeno.

Ve čtvrté části práce jsme studovali vliv mutací Val82Thr a Ile84Val HIV-1 PR na vazbu s COSANem. Malý pokles síly vazby mezi COSANem a mutovanou HIV PR je způsoben ztrátou jedné dvojevodíkové vazby.

V poslední části naší práce jsme hledali geometrie řetězce spojujícího dvě kobaltokarboránové klece GB80 v HIV PR. Nalezli jsme sérii konformerů s širokým spektrem energií, z toho pět konformerů bylo ve zvoleném rozmezí 3 kcal/mol. Těchto pět konformerů se lišilo v geometrii řetězce, zatímco centrální dusíkový atom řetězce byl vždy ve stejné pozici a odpovídal elektronové hustotě, do které byla namodelována voda 110 v krystalografické struktuře.

Závěry

i) Karborány interagují se stavebními bloky biomolekul téměř výhradně tvorbou dvojevodíkových vazeb.

- ii) Dvojvodíkové vazby jsou přednostně tvořeny vodíky v *meta* a *para* poloze.
- iii) Exo-substituce metalacarboranů způsobují vyšší narůst stabilizačních energií a nižší narůst desolvatačních energií než substituce přechodného kovu.
- iv) Tři rotamery Cb2 molekuly jsou silně populovány a dalších šest rotamerů může být přítomno ve studovaném komplexu.
- v) Cb1 molekula je více závislá na rotaci Cb2 a sedm jejích Cb1 rotamerů může být zcela vyloučeno.
- vi) Malý pokles síly vazby mezi COSANem a mutovanou HIV PR v pozicích 82 a 84 je způsoben ztrátou jedné dvojvodíkové vazby dle našeho modelu.
- vii) Nalezli jsme sérii pěti různých konformerů GB80 s centrálním dusíkovým atomem na konstatní pozici, jež odpovídá elektronové hustotě, do které byla namodelována voda číslo 110 v krystalografické struktuře.

References/ Seznam citací

A. J. Fanfrlík, M. Lepšík, D. Horinek, Z. Havlas and P. Hobza: Interaction of carboranes with biomolecules: Formation of dihydrogen bonds. *ChemPhysChem*, **2006**, 7, 1100-1105

B J. Fanfrlík, D. Hnyk, M. Lepšík and P. Hobza: Interaction of Heteroboranes with Biomolecules, Part 2: The Effect of Various Metal Vertices and Exo-substitutions. *Phys. Chem. Chem. Phys.*, **2007**, 9, 2085-2093

C J. Fanfrlík, J. Rejnek, M. Hanus and P. Hobza: Hydration Gibbs Energies of Nucleic Acid Bases Determined by Gibbs Energy Perturbation, Continuous and Hybrid Approaches. *Collect. Czech. Chem. Commun.* **2005**, 70, 1756-1768,

D J. Fanfrlík, J. Brynda, J. Řezáč, P. Hobza and M. Lepšík: Interpretation of Protein/Ligand Crystal Structure using QM/MM Calculations: Case of HIV-1 Protease/Metallacarborane complex. *J Phys. Chem. B*, submitted

E M. Kožíšek, P. Cígler, M. Lepšík, J. Fanfrlík, P. Řezáčová, J. Brynda, J. Pokorná, J. Plešek, B. Grüner, K. Grantz-Šašková, J. Václavíková, H.-G. Kraeusslich, V. Král and J. Konvalinka: Inorganic polyhedral metallacarborane inhibitors of HIV protease – a new approach to overcoming antiviral resistance. *J. Med. Chem*, under revision

F P. Řezáčová, J. Pokorná, J. Brynda, M. Kožíšek, P Cígler, M. Lepšík, J. Fanfrlík, J. Řezáč, K. Grantz-Šašková, I. Siegllová, J. Plešek, B Grüner, V Šícha, J. Sedláček, H.-G. Kraeusslich, V. Král and J. Konvalinka: Structure-based design of HIV protease inhibitors based on inorganic polyhedral metallacarboranes. *In preparation*